

TI, - Information retrieval system with speech synthesis unit in internet e.g. WWW - has speech synthesis unit that converts speech synthesis information in retrieved data to corresponding audio information which is then output by respective output unit

PR - JP19960302994 19961114

PN - JP10143352 A 19980529 DW199832 G06F3/16 006pp

PA - (NITE ) NIPPON TELEGRAPH & TELEPHONE CORP

IC - G06F3/16 ;G10L5/02 ;G10L9/00

AB - J10143352 The system includes an information retrieval unit (11) that searches information stored in a storage unit (30) and retrieves the information via a telecommunication network (20). A speech synthesis command is analysed from the retrieved information (41) and it is converted into a speech synthesis information (42) by a storage information conversion unit (12).

- An interface conversion unit (13) converts the speech synthesis information so that it corresponds to the interface of a speech synthesis unit (14). The speech synthesis unit generates an audio information (44) according to the received input. An audio information output unit (15) outputs audio information to external.

- ADVANTAGE - Converts retrieved information into audio output efficiently.

- (Dwg.1/1)

OPD - 1996-11-14

AN - 1998-367355 [32]

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平10-143352

(43)公開日 平成10年(1998) 5月29日

(51)Int.Cl.<sup>6</sup>

識別記号

F I

G 0 6 F 3/16

3 3 0

G 0 6 F 3/16

3 3 0 K

G 1 0 L 5/02

G 1 0 L 5/02

J

9/00

9/00

N

審査請求 未請求 請求項の数4 O L (全 6 頁)

(21)出願番号 特願平8-302994

(22)出願日 平成8年(1996)11月14日

(71)出願人 000004226

日本電信電話株式会社

東京都新宿区西新宿三丁目19番2号

(72)発明者 佐々木 努

東京都新宿区西新宿三丁目19番2号 日本  
電信電話株式会社内

(72)発明者 浜野 輝夫

東京都新宿区西新宿三丁目19番2号 日本  
電信電話株式会社内

(72)発明者 梶井 健

東京都新宿区西新宿三丁目19番2号 日本  
電信電話株式会社内

(74)代理人 弁理士 若林 忠

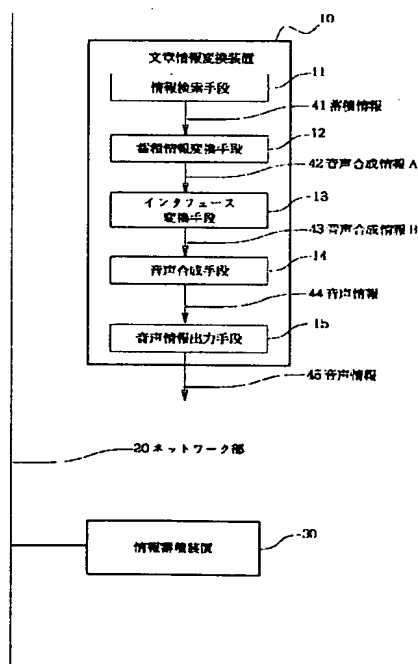
最終頁に続く

(54)【発明の名称】 文章情報変換装置および変換方法

(57)【要約】

【課題】 画像表示手段をもたない情報検索装置や画像手段を利用せずに情報検索を行なおうとする場合において、文章によつて記述された情報を効率よく音声情報に変換する文章情報変換装置と変換方法とを提供する。

【解決手段】 検索する情報が蓄積されている情報蓄積装置30から、電気通信設備などで構成されるネットワーク20を介して蓄積情報を検索する情報検索手段11と、検索された蓄積情報41から、音声合成コマンドを解析し、音声合成情報として変換する蓄積情報変換手段12と、音声合成情報42を、音声合成手段のインタフェースに合うように変換するインタフェース変換手段13と、変換された音声合成情報43から、音声情報44を生成する音声合成手段14と、音声情報45を外部に出力する音声情報出力手段15とを具備する。



## 【特許請求の範囲】

【請求項1】 蓄積された情報を効率良く音声情報に変換して提供するための文章情報変換装置であって、検索する情報が蓄積されている情報蓄積装置から、電気通信設備などで構成されるネットワークを介して蓄積情報を検索する情報検索手段と、検索された前記蓄積情報から、音声合成コマンドを解析し、音声合成情報として変換する蓄積情報変換手段と、前記音声合成情報を、音声合成手段のインタフェースに合うように変換するインタフェース変換手段と、前記変換された音声合成情報から、音声情報を生成する音声合成手段と、前記音声情報を外部に出力する音声情報出力手段と、を具備したことを特徴とする文章情報変換装置。

【請求項2】 請求項1に記載の文章情報変換装置を用いた文章情報変換方法であって、前記情報蓄積装置の音声変換の対象となる蓄積情報には、予め音声合成用コマンドが蓄積情報として加えられ、音声合成用コマンドが加えられた前記情報を、前記ネットワークを介して、前記変換装置の前記検索手段で取り込んで前記情報変換手段に送りだし、前記蓄積情報変換手段では、送られてきた前記蓄積情報から前記音声合成コマンドを解析して、音声合成情報として使用されるテキストデータと、音声合成用のコマンドとに変換して音声合成情報を生成し、生成した前記音声合成情報を前記インタフェース変換手段に送りだし、前記インタフェース変換手段では、前記音声合成情報を前記音声合成手段に対応したインタフェースに変換して前記音声合成手段に送りだし、前記音声合成手段では、前記音声合成情報からデータに従って音声合成処理を行なって音声情報を生成して、前記音声情報出力手段に送りだし、前記音声情報出力手段では、前記音声情報を外部に出力する、ことを特徴とする文章情報変換方法。

【請求項3】 前記情報蓄積手段の前記蓄積情報はHTML言語で記述され、前記音声合成用コマンドは、前記HTMLに設けられている検索システムにコメントとして理解させるためのコマンド内に記述されている、ことを特徴とする請求項2に記載の文章情報変換方法。

【請求項4】 前記情報蓄積手段の前記蓄積情報はHTML言語で記述され、前記HTMLに独自に設けたコマンドにより、検索システムに前記音声合成用コマンドによる記述があることを理解させる、ことを特徴とする請求項2に記載の文章情報変換方法。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、電気通信設備などで構成されるネットワークを介して接続された情報蓄積装置および情報検索装置間において、該情報検索装置が

所望の情報を該情報蓄積装置から検索、提示する技術分野に関し、特に音声に変換して提示する技術分野に関する。

## 【0002】

【従来の技術】インターネット上に構築されたWWW(World Wide Web)の様な、電気通信設備などで構成されるネットワークを介して接続された情報蓄積装置および情報検索装置によって構成される分散型のデータベースシステムにおいて、情報出力手段として画像表示手段がないような情報検索装置では、情報出力手段は音声情報や振動などに限定される。このような条件下では、インターネット上のWWWで使われているHTML言語の記述による文章を音声合成により音声情報に変換して出力されていた。

## 【0003】

【発明が解決しようとする課題】しかしながら、HTMLは、もともと文章の整形を前提に考案されているため、HTML文章を質のいい音声情報に変換するには、HTML文章の構造解析が必要になるという問題があった。さらに、HTML文章を解析したとしても完全に作者の意図を反映した音声情報に変換することは困難であるという問題があった。

【0004】本発明の目的は、画像表示手段をもたない情報検索装置や画像手段を利用せずに情報検索を行なおうとする場合において、文章によつて記述された情報を効率よく音声情報に変換する文章情報変換装置と変換方法とを提供することである。

## 【0005】

【課題を解決するための手段】本発明の文章情報変換装置は、蓄積された情報を効率良く音声情報に変換して提供するための文章情報変換装置であって、検索する情報が蓄積されている情報蓄積装置から、電気通信設備などで構成されるネットワークを介して蓄積情報を検索する情報検索手段と、検索された蓄積情報から、音声合成コマンドを解析し、音声合成情報として変換する蓄積情報変換手段と、音声合成情報を、音声合成手段のインタフェースに合うように変換するインタフェース変換手段と、変換された音声合成情報から、音声情報を生成する音声合成手段と、音声情報を外部に出力する音声情報出力手段とを具備する。

【0006】本発明の文章情報変換方法は、上述の文章情報変換装置を用いた文章情報変換方法であって、情報蓄積装置の音声変換の対象となる蓄積情報には、予め音声合成用コマンドが蓄積情報として加えられ、音声合成用コマンドが加えられた情報を、ネットワークを介して、変換装置の検索手段で取り込んで情報変換手段に送りだし、蓄積情報変換手段では、送られてきた蓄積情報から音声合成コマンドを解析して、音声合成情報として使用されるテキストデータと、音声合成用のコマンドとに変換して音声合成情報を生成し、生成された音声合成

情報をインタフェース変換手段に送りだし、インタフェース変換手段では、音声合成情報を音声合成手段に対応したインタフェースに変換して音声合成手段に送りだし、音声合成手段では、音声合成情報からデータに従って音声合成処理を行なって音声情報を生成して、音声情報出力手段に送りだし、音声情報出力手段では、音声情報を外部に出力する。

【0007】情報蓄積手段の蓄積情報はHTML言語で記述され、音声合成用コマンドは、HTMLに設けられている検索システムにコメントとして理解させるためのコマンド内に記述されていることが好ましい。

【0008】また、情報蓄積手段の蓄積情報はHTML言語で記述され、HTMLに独自に設けたコマンドにより、検索システムに音声合成用コマンドによる記述があることを理解させてもよい。

【0009】

【発明の実施の形態】インターネットのWWWなどで使われている文章情報を記述する言語であるHTMLには、タグと呼ばれる文章整形用のコマンドが使われている。このタグを手がかりに文章解析して音声合成情報を生成することは可能であるが、聴きやすい品質の良い音声情報へと変換するには情報不足である。これを解決す

るには、音声合成手段が理解できるようなインタフェースを介在させる必要があるものの、音声合成処理を前提としたコマンドをHTMLに包含させる必要がある。

【0010】しかしながら、HTMLは世界的に用いられる仕様であり、勝手な拡張は、音声合成手段とのインタフェースを持たない通常の検索システムなどに混乱を与えてしまうし、通常の検索システムでの閲覧を妨害してしまう。そこで、通常のHTMLを解析する検索システムに対して影響を与えないように、音声合成処理用のコマンドを埋め込む必要がある。具体的には、HTMLには、コメントとして検索システムに理解させるコマンドがあり、このコメントのコマンド行はシステムは認識していても画面などに出力されることはない。本発明では、この特性を利用し、コメント行に音声合成処理用のコマンドを記述する。したがって、音声合成処理部では、通常は無視しているコメント行に音声合成処理用のコマンドが記述されていないかをチェックすればよい。

【0011】ここで、音声合成処理用のコマンドの例を表1に示す。例えば、

【0012】

【表1】

```
<!-- \sound{speed[4], vol[8], speaker[f]} -->こんにちは、今日はいいい天気ですね。
```

とした場合は、発声スピードをレベル4、発声ボリュームをレベル8で発声、話者を女性として音声情報を出力させる。このように、表1に示すようなコマンドを、必要に応じてHTML文書中に記述することで、音声合成手段に対応していない情報検索システムに混乱を与えることなく、音声合成手段に対応した検索システムに対しては音声合成手段を直接制御することが可能となり、質の高い聴きやすい音声情報を提供することが可能となる。

【0013】また、HTML言語に、音声合成処理用のコマンドの記述がある旨を提示するタグを設けることにより、音声合成処理用のコマンドを記述してもよい。この記述方法では、まず、タグにより音声合成のための情報記述が行なわれることを示し、このタグ以降に音声合成用のテキストおよびコマンドなどの情報を記述する。この例を表2に示す。

【0014】

【表2】

```
<script language="SoundScript">
<!--
\speed[4]
\vol[8]
\speaker[f]
こんにちは、今日はいいい天気ですね。
-->
</script>
```

上記の例では、<script language="SoundScript">と</script>との間に音声合成用の情報が記述され、<!-- と -->（情報検索システムではコメントとして解されるHTML言語による記述）との間に音声合成手段を制御するコマンドと音声合成されるテキストなどの情報が記述される。この記述部分がインタフェース部を介して音声合成手段に渡され、音声出力手段により出力される。この方法を加えることにより情報検索システムの負担を小さくできる利点がある。

【0015】一方、音声合成手段および音声合成手段とのインタフェースを持っていない情報検索システムでは<script language="SoundScript">...</script>を認識できないため、この部分は無視される。

【0016】次に、本発明の実施の形態について図面を参照して説明する。図1は本発明の第1の実施の形態の

文章情報変換装置の構成を示すブロック図であり、図中符号10は文章情報変換装置、11は情報検索手段、12は蓄積情報変換手段、13はインタフェース変換手段、14は音声合成手段、15は音声情報出力手段、20はネットワーク部、30は情報蓄積装置、41は蓄積情報、42は音声合成情報A、43は音声合成情報B、44、45は音声情報である。

【0017】10の文章情報変換装置は、例えば、パーソナルコンピュータなどで構成され、11の情報検索手段や12の蓄積情報変換手段、13のインタフェース変換手段および14の音声合成手段はパーソナルコンピュータにおいて動作するアプリケーションソフトウェアなどとして実現できる。15の音声情報出力手段はスピーカ装置などである。

【0018】次に、本発明の文章情報変換方法について説明する。インターネットのWWWなどで使われているHTMLによる文章情報は、情報蓄積装置30に蓄積され、ネットワーク部20を介して、蓄積情報として検索閲覧される。本発明では、情報蓄積装置30には、通常のHTMLにより記述された蓄積情報と、HTMLと音声合成用コマンドにより記述された蓄積情報とが蓄積されることになる。ここにおける説明では、HTMLと音

声合成用コマンドにより記述された文章情報とを併せて蓄積情報とみなすことにする。

【0019】情報蓄積装置30に蓄積された蓄積情報は、情報検索手段11により、ネットワーク部20を介して文章情報変換装置10に取り込まれる。情報検索手段11では、取り込んだ蓄積情報41を必要な数量蓄積するなどしてバッファリングしながら、逐次、蓄積情報変換手段12へ送る。蓄積情報変換手段12では送られてきた蓄積情報41から音声合成用コマンドを解析し、音声合成情報として使用されるテキストデータと、音声合成時に用いられるコマンドとに変換し、音声合成情報に不要な部分は削除されて音声合成情報A42に変換される。

【0020】このとき、音声合成情報A42として解析されるコマンドとしては表3に示すようなコマンドなどがある。蓄積情報変換手段12では、表3に示すようなコマンドを蓄積情報から探し出し、表に規定されているような機能を実施できるような形式のデータに変換することになる。

【0021】

【表3】

コマンド	機能	機能説明
speed	発声速度指定	8段階
vol	発声ボリューム指定	16段階
symbol	記号読み方指定	「. ,」を読む、読まない
speaker	話者指定	男性、女性、中性（ロボット）
tone	声の高さ指定	16段階
accent	アクセント成分の大きさ（抑揚）指定	16段階
num	数字の読み方指定	桁読み、棒読み
pause	発声の一時停止	時間
silence	無変換指定	読み上げない部分の指定
comment	注釈挿入指定	注釈文を指定できる
bcsp	効果音指定	擬音や効果音の指定
lang	読み上げ言語指定	日本語、英語など

蓄積情報変換手段12で変換された音声合成情報A42は、インタフェース変換手段13で音声合成手段14が受け入れられるようなインタフェースの音声合成情報B43に変換されて音声合成手段14に送られる。音声合成手段14は送られてきたデータにしたがって音声合成処理を行い、音声情報44を音声情報出力手段15に送りだし、音声情報出力手段から外部に出力する。

【0022】このような構成になっているため、インターネットのWWWで用いられる文章情報を効率良く、品質の高い音声情報として合成することが可能となる。

【0023】

【発明の効果】本発明によって、インターネットのWW

Wで用いられる文章情報に効率が良く、品質の高い音声合成のための情報を提供することができ、効率が良く品質の高い音声合成の実現を可能とさせる効果がある。

【図面の簡単な説明】

【図1】本発明の第1の実施の形態の文章情報変換装置の構成を示すブロック図である。

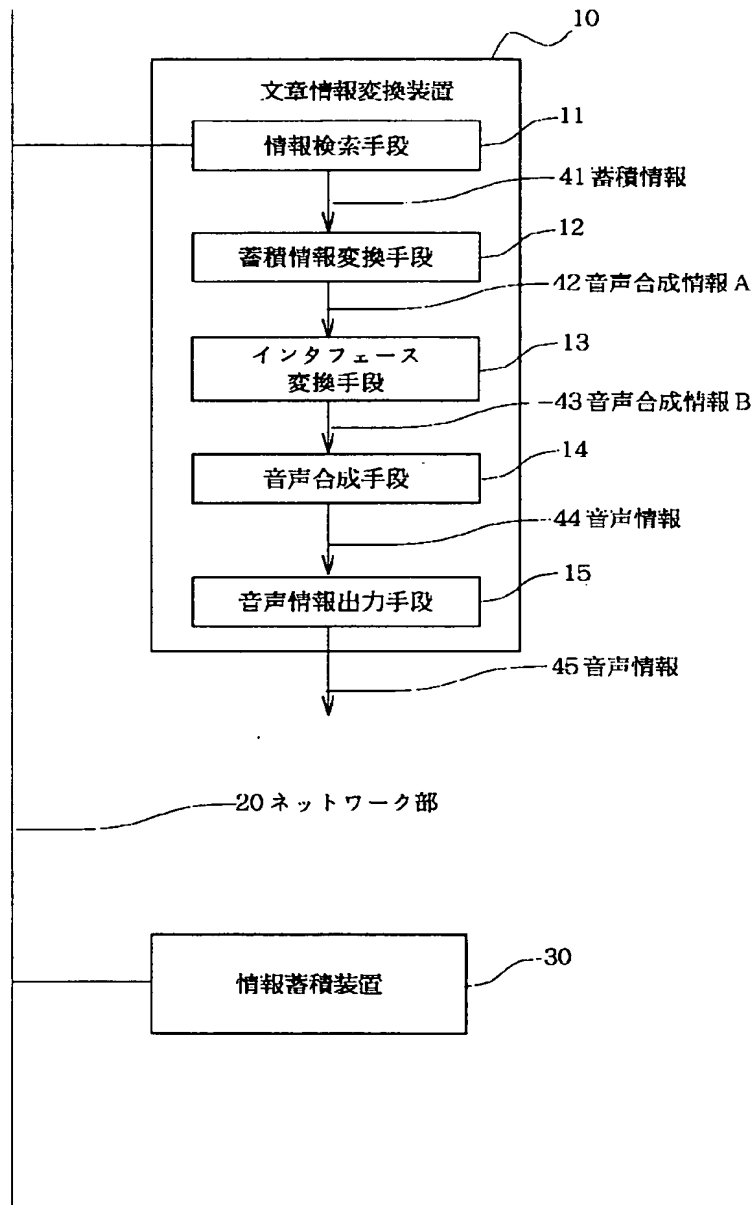
【符号の説明】

- 10 文章情報変換装置
- 11 情報検索手段
- 12 蓄積情報変換手段
- 13 インタフェース変換手段
- 14 音声合成手段

15 音声情報出力手段  
 20 ネットワーク部  
 30 情報蓄積装置  
 41 蓄積情報

42 音声合成情報A  
 43 音声合成情報B  
 44、45 音声情報

【図1】



フロントページの続き

(72)発明者 曾根原 登  
東京都新宿区西新宿三丁目19番2号 日本  
電信電話株式会社内